

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-235652

(43)Date of publication of application : 29.08.2000

(51)Int.Cl.

G06T 11/80  
 G06F 17/30  
 H04N 5/76  
 H04N 5/91  
 // G06F 3/16

(21)Application number : 11-035611

(71)Applicant : CANON INC

(22)Date of filing : 15.02.1999

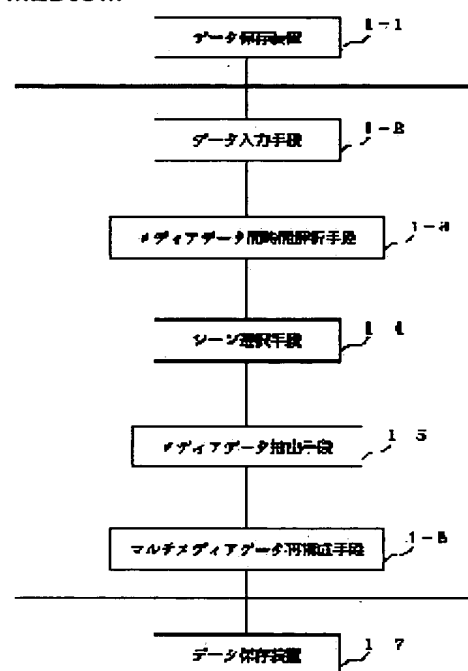
(72)Inventor : NAGAI HIROYUKI  
 SATO HIROAKI

## (54) DATA PROCESSOR, DATA PROCESSING METHOD AND MEDIUM

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To process contents formed of plural pieces of media data and to generate a digest by installing a data extraction means extracting information constituting a representative scene from respective pieces of media data and a multi-media data reconstitution means reconstituting extracted data.

**SOLUTION:** A data preservation device 1-1 is an input content and a data input means 1-2 reads contents data from the data preservation device 1-1. An inter-media time analysis means 1-3 analyzes a time relation between media data. A scene selection means 1-4 divides contents reproduction time into scenes based on an analysis result and selects a representative scene. A media data extraction means 1-5 extracts information constituting the representative scene from respective pieces of media data. A multimedia data reconstitution means 1-6 reconstitutes extracted data as the digest. The obtained digest is stored in a data preservation device 1-7.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

**THIS PAGE BLANK (USPTO)**

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開 2000-235652

(P 2000-235652A)

(43) 公開日 平成12年8月29日 (2000. 8. 29)

(51) Int. Cl. 7	識別記号	F I	テーマコード (参考)
G 0 6 T	11/80	G 0 6 F	15/62 3 2 1 A 5B050
G 0 6 F	17/30	H 0 4 N	5/76 B 5B075
H 0 4 N	5/76	G 0 6 F	3/16 3 3 0 C 5C052
	5/91		15/40 3 7 0 G 5C053
// G 0 6 F	3/16 3 3 0		15/401 3 2 0 A
審査請求 未請求 請求項の数 1 4		O L	(全 7 頁) 最終頁に続く

(21) 出願番号 特願平11-35611  
(22) 出願日 平成11年2月15日 (1999. 2. 15)

(71) 出願人 000001007  
キャノン株式会社  
東京都大田区下丸子3丁目30番2号  
(72) 発明者 長井 宏之  
東京都大田区下丸子3丁目30番2号キャノン  
株式会社内  
(72) 発明者 佐藤 宏明  
東京都大田区下丸子3丁目30番2号キャノン  
株式会社内  
(74) 代理人 100090538  
弁理士 西山 恵三 (外2名)

最終頁に続く

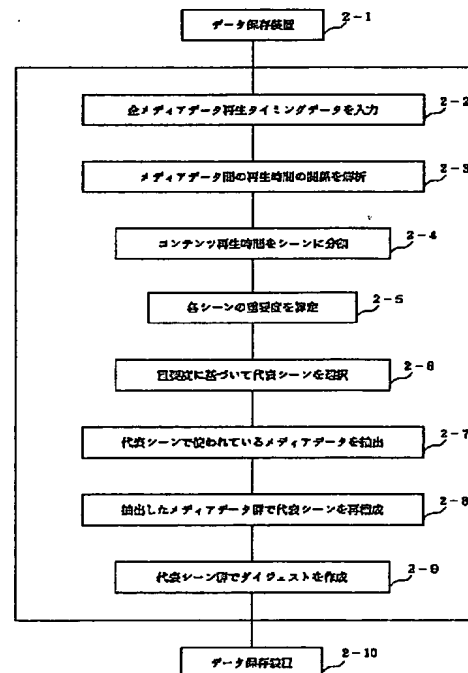
(54) 【発明の名称】 データ処理装置、データ処理方法及び媒体

(57) 【要約】

【課題】 動画、静止画、テキスト、音声といったメディアデータとそれらの再生タイミングデータからなるコンテンツからダイジェストを生成する方法を提案する。

【解決手段】 複数のメディアデータを組み合わせたコンテンツでは、各メディアデータの再生タイミングデータというものを利用することが出来る。したがって、この各メディアデータの再生タイミングを元にコンテンツのダイジェスト生成を行う方法を提案する。開示される実施形態におけるダイジェスト生成方法は次の四つのステップから成る。

1. 再生タイミングデータよりメディアデータ間の再生時間関係を解析
2. 各メディアデータの再生開始、終了時刻を用いたシーン分割
3. 各メディアデータの重要度から得るシーンの重要度算定
4. 各メディアデータからのダイジェストデータの構築



ダイジェスト生成方法流れ図

## 【特許請求の範囲】

【請求項 1】 複数のメディアデータとその時間関係を記述したデータを入力する入力手段と、メディアデータ間の時間関係を解析するメディアデータ間時間解析手段と、解析結果に基づき代表シーンを選択するシーン選択手段と、代表シーンを構成する情報を各メディアデータから抽出するメディアデータ抽出手段と、抽出したデータを再構成するマルチメディアデータ再構成手段とを有することを特徴とするデータ処理装置。

【請求項 2】 更にメディアデータのタイプに応じてその内容を解析するメディア内時間解析手段を有することを特徴とする請求項 1 のマルチメディア処理装置。

【請求項 3】 前記メディア内時間解析手段は、ビデオ映像もしくはサウンドデータの変化する点を検出する手段を含むことを特徴とする請求項 2 のマルチメディアデータ処理装置。

【請求項 4】 前記マルチメディアデータ再構成手段は HTML 形式のデータを出力とすることを特徴とする請求項 1 乃至請求項 4 のマルチメディアデータ処理装置。

【請求項 5】 前記入力手段は SMIL 形式のデータを入力とすることを特徴とする請求項 1 乃至請求項 3 のマルチメディアデータ処理装置。

【請求項 6】 前記シーン選択手段は、各メディアデータの重要度計算手段と各メディアデータの重要度を入力とするシーン全体の重要度計算手段と重要度の高いものを選択する選択手段とから構成されることを特徴とする請求項 1 のマルチメディアデータ処理装置。

【請求項 7】 複数のメディアデータとその時間関係を記述したデータを入力する入力ステップと、メディア間の時間関係を解析するメディア間時間解析ステップと解析結果に基づき代表シーンを選択するシーン選択ステップと、代表シーンを構成する情報を各メディアデータから抽出するメディアデータ抽出ステップと抽出したデータを再構成するマルチメディアデータ再構成ステップとを有することを特徴とするデータ処理方法。

【請求項 8】 更にメディアデータのタイプに応じてその内容を解析するメディア内時間解析ステップを有することを特徴とする請求項 7 のマルチメディア処理方法。

【請求項 9】 前記メディア内時間解析ステップは、ビデオ映像もしくはサウンドデータの変化する点を検出するステップを含むことを特徴とする請求項 8 のマルチメディアデータ処理方法。

【請求項 10】 該マルチメディアデータ再構成ステップは HTML 形式のデータを出力とすることを特徴とする請求項 7 乃至請求項 9 のマルチメディアデータ処理方法。

【請求項 11】 該入力ステップは SMIL 形式のデータを入力とすることを特徴とする請求項 7 乃至請求項 9 のマルチメディアデータ処理方法。

【請求項 12】 該シーン選択ステップは、各メディア

データの重要度計算ステップと各メディアデータの重要度を入力とするシーン全体の重要度計算ステップと重要度の高いものを選択する選択ステップとから構成されることを特徴とする請求項 7 のマルチメディアデータ処理方法。

【請求項 13】 複数のメディアデータとその時間関係を記述したデータを入力する入力ステップと、メディア間の時間関係を解析するメディア間時間解析ステップと解析結果に基づいた代表シーンをを用いて前記複数のメディアデータを再構成するマルチメディアデータ再構成ステップとを有することを特徴とする画像処理方法。

【請求項 14】 請求項 7 乃至請求項 12 のステップをコンピュータが実行可能な形式のプログラムとして格納するプログラム媒体。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 本発明はデータ処理装置、方法及び媒体に関し、特に、複数のメディアデータの同期再生を行うことが出来る装置で利用するコンテンツの、処理方法、装置および媒体に関するものである。

## 【0002】

【従来の技術】 近年、コンピュータの処理速度の向上により、動画と音声を用いた動的コンテンツの利用が実用に足るようになってきている。このようなコンテンツに含まれる情報を得るには、そのすべてを再生する方法があるが、概略の情報を得たい場合にはこの方法は適当ではないため、コンテンツから重要な情報をピックアップして利用者に提示するダイジェストの作成が効果的である。そこで、既に動画データからのダイジェスト作成方法については特開平 08-009310 をはじめとして多くの提案がなされている。

【0003】 またインターネットの普及により動的なコンテンツの高速な転送の必要性が出てきているが、その際動画と音声のみからなるコンテンツは、大量のデータ転送が必要になるため、文字や静止画といった比較的データ転送量の小さなメディアデータも組み合わせることによって、動的なコンテンツを生成する手段が実用化されている。

【0004】 このような複数のメディアデータを組み合わせた動的なコンテンツにおいてもダイジェストの生成は効果的であると考えられる。

## 【0005】

【発明が解決しようとする課題】 既に提案されているダイジェスト生成方法では、単一のメディアデータからなるコンテンツのダイジェスト生成には利用できるが、その他静止画やテキストも含めた複数のメディアデータからなるコンテンツのダイジェスト生成を行うことは出来ない。そこで、本出願に関わる発明の目的は複数のメディアデータからなるコンテンツを処理して例えばダイジェストを生成することにある。

## 【0006】

【課題を解決するための手段】上記目的を達成するため、本出願に係る第一の発明であるデータ処理装置は、メディアデータ間の時間関係を解析するメディアデータ間時間解析手段と、解析結果に基づき代表シーンを選択するシーン選択手段と、代表シーンを構成する情報を各メディアデータから抽出するメディアデータ抽出手段と、抽出したデータを再構成するマルチメディアデータ再構成手段とを有することを特徴とする。

## 【0007】

【発明の実施の形態】ここでは複数メディアデータの時間関係を記述したデータとしてSMIL (Synchronized Multimedia Integration Language)、(SMIL及びHTMLはW3C (World Wide Web Consortium) によるRecommendation (標準) であり、それぞれ

`http://WWW.W3C.org/TR/REC-smil`

`http://WWW.W3C.org/TR/REC-html401`

に仕様書が示されている) ダイジェストの形式としてHTML (Hyper Text Markup Language) を用いる場合について記述する。SMILデータには用いるメディアデータの存在場所、再生のタイミングが複数記述されている。HTMLデータには、テキストのダイジェストと、動画、静止画像データのダイジェスト画像データへのリンクが含まれる。本実施例はSMILデータおよび複数のメディアデータから構成されるコンテンツから、そのダイジェストとなる画像データとHTMLファイルを作成するものである。

【0008】図1に本発明の一実施形態のダイジェスト生成装置構成図を示す。図において1-1は入力コンテンツである。複数のメディアデータとそれらの再生タイミングデータであるSMILデータが格納されているデータ保存装置であり、1-2はデータ保存装置1-1からコンテンツデータを読み出すデータ入力手段、1-3はメディアデータ間の時間関係を解析するメディア間時間解析手段、1-4は解析結果に基づきコンテンツ再生時間をシーンに分割、代表シーンを選択するシーン選択手段、1-5は、代表シーンを構成する情報を各メディアデータから抽出するメディアデータ抽出手段、1-6は抽出したデータをダイジェストとして再構成するマルチメディアデータ再構成手段であり、得られたダイジェストはデータ保存装置1-7に格納される。

【0009】図2に図1に示した構成の動作、即ちダイジェスト生成方法の流れ図を示す。図2において201-1が複数のメディアデータと各メディアデータの再生タイミングデータを保有するデータ保存装置。2-2は保存装置から先のコンテンツを構成するデータ群を入力

するデータ入力ステップ、2-3はメディアデータ間の再生時間の関係を解析するメディアデータ間時間解析ステップ、2-4は解析結果に基づいてコンテンツの再生時間をシーンに分割していくシーン分割ステップ、2-5は各シーンに使われているメディアデータの重要度からシーンの重要度を算定するシーン重要度算定ステップ、2-6はシーンの重要度からダイジェストに用いるシーンを選択するシーン選択ステップ、2-7は選択された代表シーンを構成するメディアデータをシーンで使われているメディアデータから抽出するシーンデータ抽出ステップ、2-8は抽出したメディアデータ群で代表シーンを再構成する代表シーン構成ステップ、2-9は代表シーン群から出力のダイジェストデータを作成するダイジェスト生成ステップ、2-10がダイジェストデータを保存するデータ保存装置である。

【0010】本実施例をコンピュータを用いて実施した装置構成を図3に示す。図3における装置はプログラムを格納したHD装置301およびメモリ302、ネットワークと接続するネットワークI/F303、プログラムによる各種の処理を実行するCPU305、プログラム及び複数のメディアデータを媒体からロードするためのリムーバブル装置307、ディスプレイ装置310、表示用ボード312、周辺コントローラ314、キーボード316、マウス318を含む。

【0011】まずHD装置またはリムーバブル装置307に蓄積されたSMILデータを読み出す。これが図2のステップ2-2に示すデータ入力ステップである。図4は、複数のメディアデータ401とその再生タイミングデータ403の例を示す図である。このコンテンツを再生する装置は、再生タイミングデータ403を解析し、各メディアデータ401を対応する時刻に再生することによって、この複数メディアデータからなるコンテンツのタイミングのとれた再生を行うことが出来る。ここでは、理解しやすくするため再生タイミングデータ403はSMILから日本語の文章に、その表現だけを変換して示してある。これらのデータ403から、各メディア401の再生開始時刻、再生終了時刻を得る。図5は、縦軸に各メディアデータの種別、横軸に時刻を取ることによって図4で示したコンテンツでの各メディアの再生タイミングを図示したものである。図4中太線が各メディアの再生時間である。本実施例では図2のステップ2-3において全てのメディアの再生開始時刻、再生終了時刻をコンテンツ再生開始時刻が早いものから順に並べる、図中時間軸上の時刻がこれに相当する。ここまでの作業がメディアデータ間時間関係解析ステップ2-3である。

【0012】次に、得られた時刻列をシーンチェンジのタイミングとして、コンテンツの再生時間をシーンに分割する。すなわちこの場合図4の時刻7までは、時刻0-2, 2-4, 4-6, 6-7の4つのシーンに分割さ

れている。これがシーン分割ステップ 2-4 である。

【0013】次にダイジェストに用いるシーンを選択するための指標となるもののシーンの重要度を算定する。分割されたシーンの重要度はそのシーンの再生に利用される各メディアデータの重要度から求められる。各メディアの重要度の和として求める場合、次式を用いる。

$$【0014】 S_i = \sum_j s_{ij}$$

$S_i$  はシーン  $i$  の重要度、 $s_{ij}$  はシーン  $i$  で再生されるメディア  $j$  の重要度とする。

【0015】各メディアの重要度は、予め各データの例えば種類に応じて重要度として定義されているか、メディア同期データに定義されているか、そのメディアデータの持つ付加情報から求められるか、もしくは、単純に各メディアデータを同じ重み付けであってもよい。要はメディアの重要度を示すものであればよい。メディアの付加情報から求める場合には、静止画、動画、テキスト等の表示を行うデータでは例えば表示領域の大きさ、音声データでは音量の大きさと重要度とが比例すると考えてもよい。

【0016】本例では、わかり易くするため単純に各メディアデータを同じ重み付け 1 で考えるものとする。すなわち、シーン中に存在するメディアデータの総数がシーンの重要度となる。これに従い図 5 に示した各シーンの重要度を算定すると、図 6 の表に示される値になる。これがシーン重要度算定ステップ 2-5 である。

【0017】このようにして得られた各シーンの重要度に従って、重要度の高いものをピックアップすることでコンテンツのダイジェストが生成される。

【0018】図 6 に示す本例では、時刻 2-4、時刻 4-6 の、二つのシーンの重要度が高いため、この 2 つのシーンをダイジェストに利用する。

【0019】更にここから静的なダイジェストを作るために、選択したシーンの中からシーンの代表となる時刻を決定してその時刻に再生されるメディアデータを得る。本実施例ではダイジェストに HTML ファイルを利用するので、音声データはダイジェストには利用しない。シーン中に動画データが含まれる場合には、その動画データに対してダイジェスト生成方法、例えば特開平 08-009310 等に提案されている動画データからのダイジェスト生成方法を用いてシーンの代表時刻を決定する。シーン中に動画データがない場合には、そのシーン中に再生される内容はどの時刻でも同じなので（シーンはメディアデータの再生タイミングで分割しているので、シーン中にメディアの再生開始や再生終了はなく、テキスト、静止画といったデータは、シーン中同じ内容を再生する）シーン代表時刻はどの時刻を選択しても良い。ここではシーン代表時刻をそのシーン開始時刻とする。本例では、選択された二つのシーン時刻 2-4、4-6 にはどちらも動画が含まれているため、動画

のダイジェスト生成方法より各シーンの代表時刻が決定される。本例では、例えば時刻 3 と時刻 5 が各シーンの代表時刻として得られたとする。これが代表シーン選択ステップ 2-6 である。

【0020】シーン代表時刻を決定すると、ステップ 2-7 においてその時刻に再生されるダイジェストを生成する。ダイジェストは、前述の様に本実施例では HTML として生成するので、シーン中に動画データが含まれているならば、その時刻に表示されるべき画像を静止画データとして得、シーン代表画像データとする。静止画データは、シーン代表時刻に使用されるファイルをもそのままシーン代表データとして用いる。これらのシーン代表画像データのファイルのアドレスをダイジェストとして出力する HTML 中にリンクを貼っておくことで参照可能にする。テキストデータは、出力する HTML ファイル中にその内容を記述する。本例では、時刻 2-4 のシーンではシーン代表画像として、シーン代表時刻 3 に再生される。動画データ a から得られる画像データと、静止画データ b を用いる。また、時刻 4-6 のシーンではシーン代表画像として、シーン代表時刻 5 に再生される。動画データ a から得られる画像データと、静止画データ a を用いる。これが代表シーン構成ステップ 2-8 及びダイジェスト生成ステップ 2-9 である。

【0021】このようにして出来上がった HTML ファイル及びシーン代表画像データは、入力ファイルである SMIL ファイルと複数のメディアデータからなるコンテンツのダイジェストとして機能する。

【0022】本実施例においては HTML ファイル、SMIL ファイルを用いたがこれに限らず、他のファイルフォーマットであってもよい。

【0023】

【発明の効果】本出願に係る第 1 の発明によれば、複数のメディアデータとその時間関係を記述したデータを入力する入力手段と、メディアデータ間の時間関係を解析するメディアデータ間時間解析手段と、解析結果に基づき代表シーンを選択するシーン選択手段と、代表シーンを構成する情報を各メディアデータから抽出するメディアデータ抽出手段と、抽出したデータを再構成するマルチメディアデータ再構成手段とを有することにより、複数のメディアデータが存在するコンテンツの場合にも、ダイジェストの生成を行うことが出来る。

【0024】本出願に係る第 2 の発明によれば、メディアデータのタイプに応じてその内容を解析するメディア内時間解析手段を有することにより、メディアデータ内にデータの再生内容が変わるメディアデータを持つコンテンツに対してもダイジェストの生成を行うことが出来る。

【0025】本出願に係る第 3 の発明によれば、メディア内時間解析手段は、ビデオ映像もしくはサウンドデータの変化する点を検出する手段を含むことにより、メデ

メディアデータ内にデータの再生内容が変わるビデオ映像もしくはサウンドデータを持つコンテンツに対してもダイジェストの生成を行うことが出来る。

【0026】本出願に係る第4の発明によれば、マルチメディアデータ再構成手段がHTML形式のデータを出力とすることにより、動的であったコンテンツのダイジェストを静的なHTMLコンテンツとして得ることが出来る。

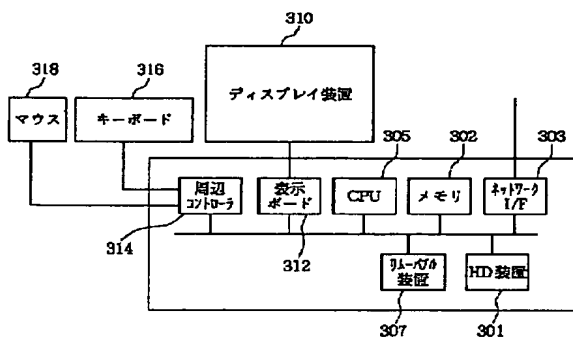
【0027】本出願に係る第5の発明によれば、入力手段がSMIL形式のデータを入力とすることにより、SMILを用いたコンテンツに対してダイジェスティングを行うことが出来る。

【0028】本出願に係る第6の発明によれば、シーン選択手段は、各メディアデータの重要度計算手段と各メディアデータの重要度を入力とするシーン全体の重要度計算手段と重要度の高いものを選択する選択手段とから構成されることにより、選択するシーンは各メディアデータから得られる情報によって決定することが出来る。

【0029】本出願に係る第7の発明によれば、複数のメディアデータとその時間関係を記述したデータを入力する入力ステップと、メディア間の時間関係を解析するメディアデータ間時間解析ステップと、解析結果に基づき代表シーンを選択するシーン選択ステップと、代表シーンを構成する情報を各メディアデータから抽出するメディアデータ抽出ステップと、抽出したデータを再構成するマルチメディアデータ再構成ステップとを有することにより、複数のメディアデータが存在するコンテンツの場合にも、ダイジェストの生成を行うことが出来る。

【0030】本出願に係る第8の発明によれば、メディアデータのタイプに応じてその内容を解析するメディア内時間解析ステップを少なくとも1つ以上有することにより、メディアデータ内にデータの再生内容が変わるメディアデータを持つコンテンツに対してもダイジェストの生成を行うことが出来る。

【図3】



本実施例をコンピュータを用いて実施した装置

【0031】本出願に係る第9の発明によれば、メディア内時間解析ステップが、ビデオ映像もしくはサウンドデータの変化する点を検出するステップを含むことにより、メディアデータ内にデータの再生内容が変わるビデオ映像もしくはサウンドデータを持つコンテンツに対してもダイジェストの生成を行うことが出来る。

【0032】本出願に係る第10の発明によれば、マルチメディアデータ再構成ステップはHTML形式のデータを出力とすることにより、動的であったコンテンツのダイジェストを静的なHTMLコンテンツとして得ることが出来る。

【0033】本出願に係る第11の発明によれば、入力ステップがSMIL形式のデータを入力とすることにより、SMILを用いたコンテンツに対してダイジェスティングを行うことが出来る。

【0034】本出願に係る第12の発明によれば、シーン選択ステップは、各メディアデータの重要度計算ステップと各メディアデータの重要度を入力とするシーン全体の重要度計算ステップと重要度の高いものを選択する選択ステップとから構成されることにより、選択するシーンは各メディアデータから得られる情報によって決定することが出来る。

【0035】本出願に係る第13、14の発明によれば、複数のメディアデータが存在するコンテンツの場合にも、ダイジェストの生成を行うことが出来る。

#### 【図面の簡単な説明】

【図1】本発明の一実施形態のダイジェスト生成装置図

【図2】ダイジェスト生成方法流れ図

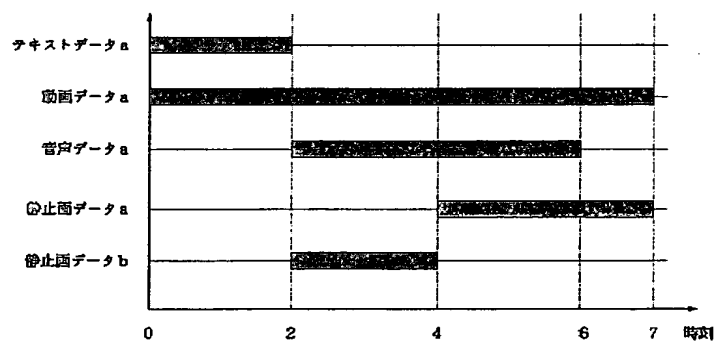
【図3】用いる装置

【図4】複数のメディアデータとその再生タイミングデータ

【図5】各メディアデータの再生タイミング

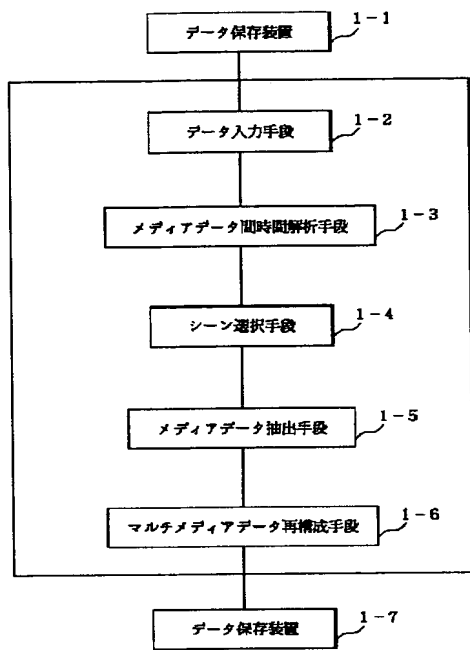
【図6】各シーンの重要度

【図5】



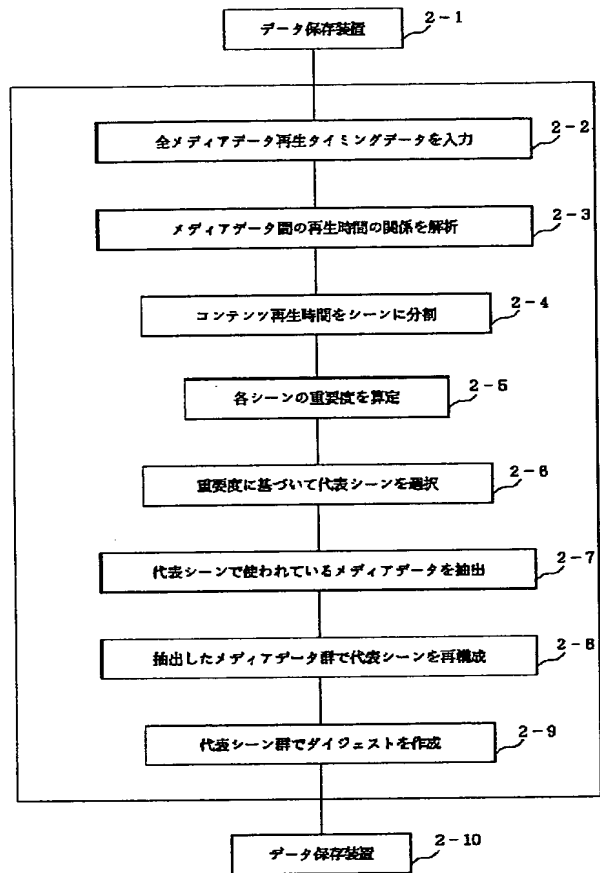
各メディアデータの再生タイミング

【図1】



ダイジェスト生成装置構成図

【図2】



ダイジェスト生成方法流れ図

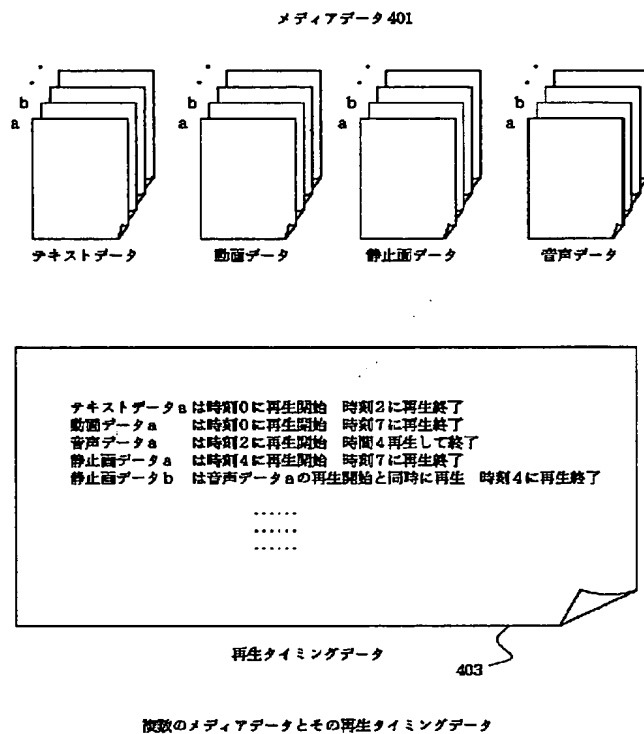
【図6】

シーン	時刻0-2	時刻2-4	時刻4-6	時刻6-7
重要度	2	3	3	2

各シーンの重要度



【図 4】



フロントページの続き

(51) Int. Cl. 7

識別記号

F I  
H O 4 N 5/91テーマコード (参考)  
N

F ターム (参考) 5B050 BA06 DA10 EA17 FA02  
 5B075 ND16 NS01 PP02 PP03 PQ02  
 PR08  
 5C052 AA01 AB04 AC08 DD02 DD04  
 DD06 EE02 EE03  
 5C053 FA14 FA23 HA29 JA01 JA16  
 JA22 KA24 LA11 LA14

**THIS PAGE BLANK (USPTO)**